

(19)日本国特許庁 ( J P )

(12) 公 開 特 許 公 報 ( A )

(11)特許出願公開番号

特開平9-120293

(43)公開日 平成9年(1997)5月6日

(51)Int.Cl. <sup>6</sup>	識別記号	片内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 5 1		G 1 0 L 3/00	5 5 1 K
	5 3 1			5 3 1 L
G 0 6 F 15/00	3 3 0		G 0 6 F 15/00	3 3 0 F

審査請求 未請求 請求項の数10 F D (全 17 頁)

(21)出願番号 特願平7-298927

(22)出願日 平成7年(1995)10月24日

(71)出願人 000006747

株式会社リコー

東京都大田区中馬込1丁目3番6号

(72)発明者 藤本 潤一郎

東京都大田区中馬込1丁目3番6号 株式  
会社リコー内

(72)発明者 宮地 達生

東京都大田区中馬込1丁目3番6号 株式  
会社リコー内

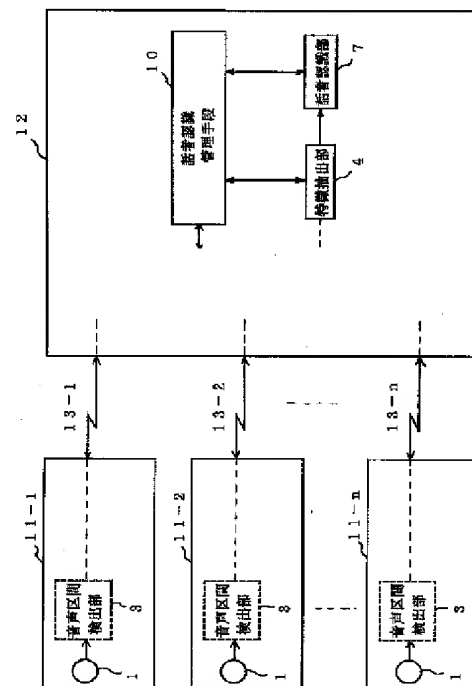
(74)代理人 弁理士 植本 雅治

(54)【発明の名称】 話者認識システムおよび話者認識方法

(57)【要約】

【課題】 利用者が例えば銀行の窓口などに出向かずとも、話者認識を行なうことができ、銀行などで提供されている入出金、残高照会などのアプリケーションを利用することが可能である。

【解決手段】 この話者認識システムは、少なくとも1つの端末11-1~11-nと中央装置12とが、情報を送受信可能に設けられている。ここで、音声入力手段1(あるいは、さらに、音声区間検出部3)が端末側に設けられ、また、特徴抽出手段4、話者認識管理手段10、話者認識部7が中央装置12に設けられたものとなっている。



**【特許請求の範囲】**

【請求項 1】 少なくとも 1 つの端末と中央装置とが、情報を送受信可能に設けられており、前記端末には、話者の音声を入力し音声信号とする音声入力手段が設けられ、また、前記中央装置には、音声信号の特徴量を抽出する特徴抽出手段と、話者認識用の情報を管理する話者認識管理手段と、前記特徴抽出手段から話者の音声の特徴量と前記話者認識管理手段によって管理されている話者認識用情報とを照合して話者認識を行なう話者認識手段とが設けられており、前記端末から前記中央装置に、音声信号を転送するようになっていないことを特徴とする話者認識システム。

【請求項 2】 少なくとも 1 つの端末と中央装置とが、情報を送受信可能に設けられており、前記端末には、話者の音声を入力し音声信号とする音声入力手段と、音声信号の特徴量を抽出する特徴抽出手段とが設けられ、また、前記中央装置には、話者認識用の情報を管理する話者認識管理手段と、話者の音声の特徴量と該話者認識管理手段によって管理されている話者認識用情報とを照合して話者認識を行なう話者認識手段とが設けられており、前記端末から中央装置に、話者の音声の特徴量を転送するようになっていないことを特徴とする話者認識システム。

【請求項 3】 請求項 2 記載の話者認識システムにおいて、さらに、前記中央装置から端末には、所定の情報が転送され、前記端末に設けられている特徴抽出手段は、中央装置から提供された情報に基づいて、入力された音声の特徴量に変換することを特徴とする話者認識システム。

【請求項 4】 請求項 1 または請求項 2 記載の話者認識システムにおいて、前記端末には、音声入力手段からの音声信号または特徴抽出手段からの特徴量を所定の暗号情報に基づいて暗号化する暗号化手段がさらに設けられており、前記端末は、音声入力手段からの音声信号または特徴抽出手段からの特徴量を所定の暗号情報に基づいて暗号化して、前記中央装置に転送することを特徴とする話者認識システム。

【請求項 5】 請求項 4 記載の話者認識システムにおいて、前記中央装置には、暗号情報を生成する暗号情報生成手段と、復号化手段とが設けられており、前記端末の暗号化手段は、前記中央装置の暗号情報生成手段によって生成された暗号情報が中央装置から提供されるとき、該暗号情報に基づいて、音声入力手段からの音声信号または特徴抽出手段からの特徴量に暗号化を施し、また前記中央装置では、端末から暗号化された音声信号または特徴量が転送されるとき、該暗号化された音声信号または特徴量を前記暗号情報に基づいて復号化手段によって復号化するようになっていないことを特徴とする話者認識システム。

【請求項 6】 請求項 4 または請求項 5 記載の話者認識

システムにおいて、前記暗号情報は、前記端末から前記中央装置への通信毎に、異なるものが用いられることを特徴とする話者認識システム。

【請求項 7】 少なくとも 1 つの端末と中央装置とが、情報を送受信可能に設けられており、前記端末に話者が音声を入力すると、話者の音声信号は中央装置に転送され、前記中央装置では、前記端末から転送された話者の音声信号から特徴量を抽出し、該特徴量と該中央装置に予め管理されている話者認識用情報とを照合して話者認識を行なうことを特徴とする話者認識方法。

【請求項 8】 少なくとも 1 つの端末と中央装置とが、情報を送受信可能に設けられており、前記端末に話者が音声を入力すると、話者の音声信号は特徴量に変換されて中央装置に転送され、前記中央装置では、前記端末から転送された話者の音声の特徴量と該中央装置に予め管理されている話者認識用情報とを照合して話者認識を行なうことを特徴とする話者認識方法。

【請求項 9】 請求項 8 記載の話者認識方法において、前記端末において特徴量の変換処理がなされるに先立って、前記中央装置から端末には所定の情報が転送され、前記端末は、中央装置から提供された情報に基づいて、入力された音声信号を特徴量に変換することを特徴とする話者認識方法。

【請求項 10】 請求項 7 または請求項 8 記載の話者認識方法において、前記端末は、さらに、音声信号または特徴量に所定の暗号情報に基づいて暗号化を施して、前記中央装置に転送することを特徴とする話者認識方法。

**【発明の詳細な説明】****【0001】**

【発明の属する技術分野】 本発明は、話者認識を行なう話者認識システムおよび話者認識方法に関する。

**【0002】**

【従来の技術】 従来、銀行などにおいて、例えば入出金、残高照会などのアプリケーションを利用者が利用する際、この利用者が正規の利用者本人であることを確認するために、暗証番号などを利用者に入力させるようにしている。また、コンピュータでは、パスワードと称して、暗証番号と同様の暗証文字列を利用者に入力させることによって本人の確認を行なっている。しかしながら、このような暗証番号や暗証文字列などの入力による確認は、他人が、暗証番号や暗証文字列を知りさえすれば、難無く、これを盗用することができる。しかも、暗証番号や暗証文字列は、それを登録した者(本人)の生年月日や記念日、あるいは電話番号、氏名の綴りなどを利用したものが多く、他人がこれを見破ることは差程難しいことではない。

【0003】 暗証番号や暗証文字列のこのような欠点を回避するため、近年、声によって本人か否かを判定する、いわゆる話者認識が着目されている。この話者認識は、ある話者が発声した音声の特徴量(特徴パターン)

が、予め登録されているこの話者の音声標準パターンと一致するか否かを調べることにより、本人か否かを判定（認識）するものである。すなわち、話者の音声から抽出した特徴量（特徴パターン）とこの話者の音声標準パターンとの類似度を計算し、類似度の高低によって本人か否かを判定するものであり、人間の肉体的特徴を利用するものであることから、音声は、暗証番号や暗証文字列に比べて他人がこれを真似ることは難しく、従って、他人の盗用をより有効に防止することができる。

【0004】

【発明が解決しようとする課題】しかしながら、このような話者認識システムは、従来一般的に、例えば銀行の窓口などに設置され、従って、話者認識を行なうためには、利用者は、その都度、銀行の窓口等へ出向かなければならないという問題があった。

【0005】本発明は、利用者が例えば銀行の窓口などに出向かずとも、話者認識を行なうことができ、銀行などで提供されている入出金、残高照会などのアプリケーションを利用することの可能な話者認識システムおよび話者認識方法を提供することを目的としている。

【0006】

【課題を解決するための手段】上記目的を達成するために、請求項1記載の発明は、少なくとも1つの端末と中央装置とが、情報を送受信可能に設けられており、端末には、話者の音声を入力し音声信号とする音声入力手段が設けられ、また、中央装置には、音声信号の特徴量を抽出する特徴抽出手段と、話者認識用の情報を管理する話者認識管理手段と、特徴抽出手段から話者の音声の特徴量と話者認識管理手段によって管理されている話者認識用情報とを照合して話者認識を行なう話者認識手段とが設けられており、端末から中央装置に、音声信号を転送するようになっていないことを特徴としている。

【0007】また、請求項2記載の発明は、少なくとも1つの端末と中央装置とが、情報を送受信可能に設けられており、端末には、話者の音声を入力し音声信号とする音声入力手段と、音声信号の特徴量を抽出する特徴抽出手段とが設けられ、また、中央装置には、話者認識用の情報を管理する話者認識管理手段と、話者の音声の特徴量と該話者認識管理手段によって管理されている話者認識用情報とを照合して話者認識を行なう話者認識手段とが設けられており、端末から中央装置に、話者の音声の特徴量を転送するようになっていないことを特徴としている。

【0008】また、請求項3記載の発明は、請求項2記載の話者認識システムにおいて、さらに、中央装置から端末には、所定の情報が転送され、端末に設けられている特徴抽出手段は、中央装置から提供された情報に基づいて、入力された音声の特徴量に変換することを特徴としている。

【0009】また、請求項4記載の発明は、請求項1ま

たは請求項2記載の話者認識システムにおいて、端末には、音声入力手段からの音声信号または特徴抽出手段からの特徴量を所定の暗号情報に基づいて暗号化する暗号化手段がさらに設けられており、端末は、音声入力手段からの音声信号または特徴抽出手段からの特徴量を所定の暗号情報に基づいて暗号化して、中央装置に転送することを特徴としている。

【0010】また、請求項5記載の発明は、請求項4記載の話者認識システムにおいて、中央装置には、暗号情報を生成する暗号情報生成手段と、復号化手段とが設けられており、端末の暗号化手段は、中央装置の暗号情報生成手段によって生成された暗号情報が中央装置から提供されるとき、該暗号情報に基づいて、音声入力手段からの音声信号または特徴抽出手段からの特徴量を暗号化を施し、また中央装置では、端末から暗号化された音声信号または特徴量が転送されるとき、該暗号化された音声信号または特徴量を暗号情報に基づいて復号化手段によって復号化するようになっていることを特徴としている。

【0011】また、請求項6記載の発明は、請求項4または請求項5記載の話者認識システムにおいて、暗号情報は、端末から中央装置への通信毎に、異なるものが用いられることを特徴としている。

【0012】また、請求項7記載の発明は、少なくとも1つの端末と中央装置とが、情報を送受信可能に設けられており、端末に話者が音声を入力すると、話者の音声信号は中央装置に転送され、中央装置では、端末から転送された話者の音声信号から特徴量を抽出し、該特徴量と該中央装置に予め管理されている話者認識用情報とを照合して話者認識を行なうことを特徴としている。

【0013】また、請求項8記載の発明は、少なくとも1つの端末と中央装置とが、情報を送受信可能に設けられており、端末に話者が音声を入力すると、話者の音声信号は特徴量に変換されて中央装置に転送され、中央装置では、端末から転送された話者の音声の特徴量と該中央装置に予め管理されている話者認識用情報とを照合して話者認識を行なうことを特徴としている。

【0014】また、請求項9記載の発明は、請求項8記載の話者認識方法において、端末において特徴量の変換処理がなされるに先立って、中央装置から端末には所定の情報が転送され、端末は、中央装置から提供された情報に基づいて、入力された音声信号を特徴量に変換することを特徴としている。

【0015】また、請求項10記載の発明は、請求項7または請求項8記載の話者認識方法において、端末は、さらに、音声信号または特徴量に所定の暗号情報に基づいて暗号化を施して、中央装置に転送することを特徴としている。

【0016】

【発明の実施の形態】図1は一般的な話者認識システム

の構成例を示す図である。図 1 を参照すると、この話者認識システムは、例えば銀行などにおける本人の確認を話者認識により行なうためのものであって、利用者の音声を入力するための音声入力手段(例えば、マイクロフォン) 1 と、利用者に所定の指定情報を入力させるための指定手段(例えばキーボード) 2 と、音声入力手段 1 から入力された信号の中から話者の音声の部分のみを音声区間として検出する音声区間検出部 3 と、音声区間検出部 3 で検出した音声区間内の音声信号から特徴量(特徴パターン)を抽出する特徴抽出部 4 と、話者認識を行なうに先立って話者の音声の標準的な特徴量(特徴パターン)を標準パターンとして話者認識用情報記憶部 5 に予め登録する登録部 6 と、利用者(話者)の音声の特徴量(特徴パターン)と話者認識用情報記憶部 5 に登録されている標準パターンとを照合し、その類似度に基づいて話者認識を行なう話者認識部 7 と、標準パターンの登録を行なう登録モードと話者認識を行なう認識モードとの切替を行なう切替部(例えばスイッチ) 8 とを有している。

【0017】ここで、特徴抽出部 4 は、音声信号を特徴量(特徴パターン)として、スペクトルに変換しても良いし、あるいは L P C ケプストラムに変換しても良く、特徴量の種類については特に限定するものではない。なお、スペクトルに変換するためには、特徴量変換には F F T を用い、また、L P C ケプストラムに変換するためには L P C 分析などを用いるのがよい。

【0018】また、標準パターンの登録時(登録モード時)において、登録部 6 は、ある話者が発声した音声に基づいて特徴抽出部 4 で抽出された特徴量(特徴パターン)を標準パターンとして話者認識用情報記憶部 5 に登録する際、図 2 に示すように、この話者により指定手段 2 から入力された指定情報(例えば、この話者の名前や生年月日、あるいはこの話者の暗証番号など)と対応付けて、標準パターンを話者認識用情報記憶部 5 に登録することができる。換言すれば、話者認識用情報記憶部 5 には、話者認識に必要な話者認識用の情報が登録されるようになっており、また、この話者認識用情報記憶部 5 には、複数の話者(例えば利用者 A, B, C, D, …)の話者認識用情報が登録可能となっている。

【0019】また、話者認識用情報記憶部 5 に登録される音声の標準パターンとしては、この話者認識システムの使用形態等に応じて、各利用者(話者)に予め言葉を発声させたものであっても良いし、各利用者ごとにそれぞれ自由に所望の言葉を発声させたものであっても良い。

【0020】また、話者認識部 7 は、例えば、古井著「デジタル音声処理」(東海出版会)などに記載されているように、現在の話者の音声の特徴パターンが話者認識用情報記憶部 5 に登録されている複数の話者の標準パターンのうちのどれに最も類似しているかを判定し、登録されている複数の話者のうちから 1 人の話者を識別する話者識別方式のものであっても良いし、話者認識用情

報記憶部 5 に登録されている複数の話者の標準パターンから現在の話者に対応する標準パターンを取り出し、この標準パターンと現在の話者の特徴パターンとを照合し、その類似度が所定基準値(しきい値)よりも高いか低いかにより現在の話者が正規の話者本人であるか否かを判定する話者照合方式のものであっても良い。

【0021】さらに、話者認識部 7 は、話者認識用情報記憶部 5 に登録される音声の標準パターンが各利用者(話者)に予め言葉を発声させたものである場合には、これに対応した認識を行なうものにすることができ、また、話者認識用情報記憶部 5 に登録される音声の標準パターンが各利用者ごとにそれぞれ自由に所望の言葉を発声させたものである場合には、これに対応した認識を行なうものにすることができる。但し、各利用者(話者)に予め決められた言葉を発声させて話者認識を行なう場合、類似の判定基準(しきい値)を各話者に対して全て一定値にすることができ、各利用者ごとにそれぞれ所望の言葉を発声させて話者認識を行なう場合には、類似の判定基準(しきい値)を各話者ごとに相違させることもできる。

【0022】以下では、説明の便宜上、話者認識システムは、各利用者(話者)に予め決められた言葉(特定の言葉)を発声させるものとし、また、話者認識部 7 では、話者照合方式の話者認識がなされるとする。なお、話者認識部 7 において、話者照合方式の話者認識がなされる場合、この話者認識時に、利用者(話者)は、指定手段 2 から登録モード時に入力した指定情報と同じ指定情報を入力する必要がある。これにより、話者認識部 7 では、話者認識用情報記憶部 5 に登録されている複数の話者の標準パターンのうちから現在の話者に対応する標準パターンを取り出すことができ、この標準パターンと現在の話者の音声の特徴パターンとの照合を行なうことができる。

【0023】このような構成の話者認識システムを利用者(例えば D)が始めて利用する場合、この利用者(話者) D は、まず、自己の音声を標準パターンとして登録する必要がある。このため、この利用者 D は、切替部(例えばスイッチ) 8 を操作して、特徴抽出部 4 を登録部 6 に接続し、登録モードに設定する。

【0024】次いで、利用者(話者) D は、指定手段 2 から所定の指定情報、例えば(利用者 D)を入力する。また、この際、利用者は、予め決められた特定の言葉を発声する。この音声は、音声入力手段 1 から入力し、音声区間検出部 3、特徴抽出部 4 により、特徴量(特徴パターン)に変換され、この話者の音声の標準パターンとして、登録部 6 に与えられる。

【0025】これにより、登録部 6 は、この利用者(話者) D の音声の標準パターンを指定手段 2 から入力された指定情報と対応付けて、話者認識用情報記憶部 5 に登録する。例えば過去に、この話者認識用情報記憶部 5 に

複数の利用者(異なる利用者) A, B, C が自己の音声を標準パターンとして登録しており、現在の利用者 D が上記のように自己の音声を標準パターンとして登録するとき、この標準パターンは、話者認識用情報記憶部 5 に図 2 に示すように記憶(登録)される。

【0026】このようにして、この音声の標準パターンが話者認識用情報記憶部 5 に記憶されると、利用者 D は、この話者認識システムにより、利用者 D についての話者認識を行なわせることができる。すなわち、この利用者 D は、このシステムを用いて、いま利用している利用者が利用者 D 本人であるか否かの判定を行なわせることができる。

【0027】具体的に、利用者 D が以後、このシステムを利用する場合、利用者 D は、切替部 8 を操作して、特徴抽出部 4 を話者認識部 7 に接続し、このシステムを認識モードに設定する。

【0028】次いで、利用者 D は、指定手段 2 から所定の指定情報、例えば(利用者 D)を入力する。また、この際、利用者 D は、予め決められた特定の言葉を発声する。この音声は、音声入力手段 1 から入力し、音声区間検出部 3、特徴抽出部 4 により、特徴量(特徴パターン)に変換されて、話者認識部 7 に与えられる。

【0029】これにより、話者認識部 7 は、指定手段 2 から入力された指定情報(利用者 D)に対応させて登録されている標準パターンを話者認識用情報記憶部 5 から取り出し、この標準パターンと特徴抽出部 4 からの特徴パターンとを照合して、その類似度を算出し、この類似度が所定基準値よりも高いか低いかを判定する。この結果、類似度が低いと判定されたときには、利用者が正規の話者本人 D ではないと判別し、この利用者による利用を拒絶する。これに対し、類似度が高いと判定されたときには、利用者が正規の話者本人 D であると判別し、利用者による利用を許可する。すなわち、利用者によるアプリケーション(例えば入出金、残高照会などの処理)の利用を許可する。

【0030】ところで、図 1 のような話者認識システムは、従来一般的に、例えば銀行の窓口などに設置され、従って、話者認識を行なうためには、利用者は、その都度、銀行の窓口等へ出向かなければならないという問題があった。

【0031】本発明は、このような問題を回避し、利用者が、利用者の自宅において、あるいは利用者の会社等において、話者認識を行ない、銀行などのアプリケーション(入出金、残高照会などのアプリケーション)等を利用できるようにすることを意図している。

【0032】図 3 は本発明に係る話者認識システムの第 1 の構成例を示す図である。この第 1 の構成例では、話者認識システムは、少なくとも 1 つの端末 11-1 ~ 11-n と中央装置 12 とが、情報を送受信可能に設けられている(例えば有線あるいは無線の通信手段 13-1

~ 13-n によって通信可能に設けられている)。ここで、各端末 11-1 ~ 11-n は、説明の便宜上、同じ構成のものであるとする。

【0033】図 3 の構成例では、図 1 の構成例において、音声入力手段 1 (あるいは、さらに、音声区間検出部 3) が端末側に設けられ、また、特徴抽出手段 4、話者認識管理手段 10、話者認識部 7 が中央装置 12 に設けられたものとなっている。なお、ここで、話者認識管理手段 10 は、図 1 の話者認識用情報記憶部 5、登録部 6 の機能を有し、さらに、これに話者認識全体の管理、制御機能、さらには、切替部 8 の機能をももたせることもできる。すなわち、ここでは、話者認識管理手段 10 とは、話者認識に関する管理、制御する機能を総称した包括的な手段であるとする。

【0034】図 4 は図 3 の話者認識システムの具体例を示す図である。なお、図 4 では、簡単のため、1 つの端末 11-1 だけが図示されているが、他の端末 11-2 ~ 11-n も、端末 11-1 と同様の構成のものであるとする。図 4 を参照すると、端末 11-1 には、音声入力手段 1、指定手段 2、音声区間検出部 3 が設けられ、さらに、端末 11-1 には、中央装置 12 との間で情報を通信手段(例えば電話回線あるいは無線) 13-1 を介して送受信するための送受信インタフェース部 14-1 が設けられている。また、中央装置 12 には、特徴抽出部 4、話者認識用情報記憶部 5、登録部 6、話者認識部 7、切替部 8 が設けられ、さらに中央装置 12 には、各端末 11-1 ~ 11-n との間で情報を送受信するための送受信インタフェース部 15 がさらに設けられている。

【0035】ここで、各端末 11-1 ~ 11-n としては、例えば電話機やパソコン(マイクロフォン、A/D 変換などの音声取込機能を備えたパソコン)を用いることが可能であって、各端末 11-1 ~ 11-n の利用者は、自己の端末を、例えば自宅や会社において、保有することができる。

【0036】また、図 4 の構成例において、中央装置 12 の送受信インタフェース 15 には、例えば交換器を用いることができる。また、特徴抽出部 4、登録部 6、話者認識部 7 は、この中央装置に搭載される特徴抽出ソフト、登録ソフト、話者認識ソフト(話者認識プログラム)として実現され、話者認識用情報記憶部 5 には、中央装置 12 に設けられているメモリを用いることができる。また、切替部 8 は、例えば、端末からの指示が標準パターンの登録(あるいは変更)であるか、話者認識であるかに応じて、登録ソフトを実行するか、話者認識ソフトを実行するかを選択する選択プログラムとして構成できる。

【0037】また、図 4 の構成例において、話者認識用情報記憶部 5 には、図 5 に示すように、標準パターンの他に、標準パターンの更新日時、類似度のしきい値など

の各種の付随情報が記憶されても良い。この場合、例えば、話者認識部 7 は、ある話者の音声の特徴パターンを、話者認識用情報記憶部 5 に記憶されている標準パターンと照合して、この話者の特徴パターンと標準パターンとの類似度を求めたとき、この標準パターンに対応するしきい値(付随情報)を話者認識用情報記憶部 5 から読み出し、上記類似度がこのしきい値よりも高いか低いかにより、この話者が本人であるか否かを判別することができる。

【0038】このような第 1 の構成例の話者認識システムでは、標準パターンの登録(さらには標準パターンの変更あるいは更新)や話者認識を行なうために、利用者は、利用者の家庭や会社等に設置されている端末を操作することによって、例えば銀行の窓口などに設置されている中央装置(例えば話者認識装置ユニット)に対し、標準パターンの登録操作、話者認識操作を、前述したと同様にして行なうことができる。

【0039】例えば、話者認識を行なうとき、利用者が、自己の端末、例えば 11-1 の指定手段 2 から、例えば、話者認識を行なう旨の指示を入力すると、この指定情報は、通信手段 13-1 を介して中央装置 12 に伝えられ、中央装置 12 では、切替部 8 を話者認識部 7 側に切替設定する。次いで、この利用者が、指定手段 2 から所定の指定情報、例えば(利用者 D)を入力すると、この指定情報は、通信手段 13-1 を介して中央装置 12 に伝えられる。また、この際、利用者 D は、予め決められた特定の言葉を発声する。この音声は、音声入力手段 1 から入力し、例えば音声区間検出部 3 から音声信号として出力され、通信手段 13-1 を介して中央装置 12 に伝えられる。中央装置 12 では、伝送された音声信号を特徴抽出部 4 により、特徴量(特徴パターン)に変換し、話者認識部 7 に与える。

【0040】これにより、話者認識部 7 は、指定手段 2 から伝送された指定情報(利用者 D)に対応させて登録されている標準パターンを話者認識用情報記憶部 5 から取り出し、この標準パターンと特徴抽出部 4 からの特徴パターンとを照合して、その類似度を算出し、この類似度が所定基準値(しきい値)よりも高いか低いかを判定する。この結果、類似度が低いと判定されたときには、利用者が正規の話者本人 D ではないと判別し、この利用者による利用を拒絶する。これに対し、類似度が高いと判定されたときには、利用者が正規の話者本人 D であると判別し、利用者による利用を許可する。すなわち、利用者によるアプリケーション(例えば入出金、残高照会などの処理)の利用を許可する。

【0041】このように、第 1 の構成例の話者認識システムでは、利用者の自宅あるいは会社等に設置されている端末(例えば電話装置)を用いて、銀行などに設置されている中央装置(例えば話者認識装置ユニット)にアクセスして、話者認識を行なわせ、話者認識の結果、本人で

あることが確認された後、入出金、残高照会などのアプリケーションを利用することができる。すなわち、利用者は、銀行等にその都度出向かずとも、自宅や会社などから銀行等の話者認識装置ユニットに話者認識を行なわせ、銀行等のアプリケーションを利用することができる。

【0042】特に、この第 1 の構成例では、利用者側の端末としては、単に、音声入力手段 1 (あるいは、これとともに指定手段 2)だけが備わったものであっても良く、従って、端末として、通常の電話器、携帯電話を用いることができる。すなわち、電話器のハンドセットの送話器を音声入力手段 1 として用い、また、電話器の操作部(プッシュホン電話器のテンキー部)を指定手段 2 として用いることができる。

【0043】但し、この第 1 の構成例では、音声入力手段 1 (あるいは音声区間検出部 3)からの音声信号を通信手段(例えば電話回線や無線)を介して中央装置 12 に送信するため、通信手段(電話回線や無線)の品質や通信環境などが悪い場合には、通信手段を伝送中に、音声信号に雑音が混入したり、音声信号が劣化し、中央装置 12 において話者認識を正しく行なうことができず、本人であるにもかかわらず、本人でないと判別されてアプリケーションを利用できないといった事態が考えられる。また、音声信号には、これをデジタル信号に変換してもデータ量が多く、従って、音声信号を通信手段(電話回線)を介して中央装置に送信する場合、相当の伝送時間を要してしまうという問題も考えられる。

【0044】図 6 は本発明に係る話者認識システムの第 2 の構成例を示す図であり、この第 2 の構成例では、第 1 の構成例における上記のような問題を回避することを意図している。すなわち、この第 2 の構成例においても、少なくとも 1 つの端末 31-1 ~ 31-n と中央装置 32 とが、情報を送受信可能に設けられている(例えば有線あるいは無線の通信手段 33-1 ~ 33-n によって通信可能に設けられている)が、この第 2 の構成例では、音声入力手段 1 (さらには音声区間検出部 3)、特徴抽出部 4 が端末側に設けられ、また、話者認識管理手段 10、話者認識部 7 が中央装置 32 に設けられている。

【0045】図 7 は図 6 の話者認識システムの具体例を示す図である。なお、図 7 では、簡単のため、1 つの端末 31-1 だけが図示されているが、他の端末 31-2 ~ 31-n も、端末 31-1 と同様の構成のものであるとする。図 7 を参照すると、端末 31-1 には、音声入力手段 1、指定手段 2、音声区間検出部 3、特徴抽出部 4 が設けられ、さらに、端末 31-1 には、中央装置 32 との間で情報を通信手段 33-1 を介して送受信するための送受信インタフェース部 34-1 が設けられている。また、中央装置 32 には、話者認識用情報記憶部 5、登録部 6、話者認識部 7、切替部 8 が設けられ、さ

らに、中央装置32には、各端末31-1～31-nとの間で情報を送受信するための送受信インタフェース部35が設けられている。

【0046】ここで、各端末31-1～31-nとしては、例えばパソコン(マイクロフォン、A/D変換などの音声取込機能を備えたパソコン)を用いることが可能であって、各端末31-1～31-nの利用者は、自己の端末を、例えば自宅や会社において、保有することができる。より具体的に、各端末31-1～31-nには、既存のパソコン(パソコン通信機能を備えたパソコン)を用いることができ、この場合、図7の構成例において、端末31-1の送受信インタフェース部34-1は、例えば、パソコンに内蔵されているモデムとして実現され、また、音声入力手段1は、パソコンに設けられているマイクロフォンで実現され、指定手段2はパソコンのコンソールで実現され、音声区間検出部3、特徴抽出部4は、パソコンに搭載されるソフトウェア、例えば、音声区間検出ソフト、特徴抽出ソフトとして実現される。

【0047】また、図7の構成例において、中央装置32の送受信インタフェース35には、例えば交換器を用いることができる。また、登録部6、話者認識部7は、この中央装置に搭載される登録ソフト、話者認識ソフト(話者認識プログラム)として実現され、話者認識用情報記憶部5には、中央装置32に設けられているメモリを用いることができる。また、切替部8は、例えば、端末からの指示が標準パターンの登録(あるいは変更)であるか、話者認識であるかに応じて、登録ソフトを実行するか、話者認識ソフトを実行するかを選択する選択プログラムとして構成できる。

【0048】また、図7の構成例においても、図4の構成例と同様、話者認識用情報記憶部5には、図5に示したように、標準パターンの他に、標準パターンの更新日時、類似度のしきい値などの各種の付随情報が記憶されても良い。この場合、例えば、話者認識部7は、ある話者の音声の特徴パターンを、話者認識用情報記憶部5に記憶されている標準パターンと照合して、この話者の特徴パターンと標準パターンとの類似度を求めたとき、この標準パターンに対応するしきい値(付随情報)を話者認識用情報記憶部5から読み出し、上記類似度がこのしきい値よりも高いか低いにより、この話者が本人であるか否かを判別することができる。

【0049】このような第2の構成例の話者認識システムにおいても、標準パターンの登録(さらには標準パターンの変更あるいは更新)、話者認識を行なうために、利用者は、利用者の家庭や会社等に設置されている端末を操作することによって、例えば銀行の窓口などに設置されている話者認識装置ユニットに対し、標準パターンの登録操作、話者認識操作を、前述したと同様に行なうことができる。

【0050】例えば、話者認識を行なうとき、利用者は、自己の端末、例えば31-1の指定手段2から、例えば、話者認識を行なう旨の指示を入力すると、この指定情報は、通信手段33-1を介して中央装置32に伝えられ、中央装置32では、切替部8を話者認識部7側に切替設定する。次いで、この利用者が、指定手段2から所定の指定情報、例えば(利用者D)を入力すると、この指定情報は、通信手段33-1を介して中央装置32に伝えられる。また、この際、利用者Dは、予め決められた特定の言葉を発声する。この音声は、音声入力手段1から入力し、例えば音声区間検出部3から音声信号として出力され、特徴抽出部4で特徴量に変換されて、通信手段33-1を介して中央装置32に伝えられる。中央装置32では、伝送された特徴量(特徴パターン)信号を話者認識部7に与える。

【0051】これにより、話者認識部7は、指定手段2から伝送された指定情報(利用者D)に対応させて登録されている標準パターンを話者認識用情報記憶部5から取り出し、この標準パターンと特徴量(特徴パターン)とを照合して、その類似度を算出し、この類似度が所定基準値よりも高いか低いかを判定する。この結果、類似度が低いと判定されたときには、利用者が正規の話者本人Dではないと判別し、この利用者による利用を拒絶する。これに対し、類似度が高いと判定されたときには、利用者が正規の話者本人Dであると判別し、利用者による利用を許可する。すなわち、利用者によるアプリケーション(例えば入出金、残高照会などの処理)の利用を許可する。

【0052】このように、第2の構成例の話者認識システムにおいても、利用者の自宅あるいは会社等に設置されている端末(例えばパソコン)を用いて、銀行などに設置されている中央装置(例えば話者認識装置ユニット)にアクセスして、話者認識を行なわせ、話者認識の結果、本人であることが確認された後、入出金、残高照会などのアプリケーションを利用することができる。すなわち、利用者は、銀行等にその都度出向かずとも、自宅や会社などから銀行等の話者認識装置ユニットに話者認識を行なわせ、銀行等のアプリケーションを利用することができる。

【0053】また、この第2の構成例では、利用者側の端末として、既存のパソコン(パソコン通信機能を備えたパソコン)を用いることができる。

【0054】さらに、この第2の構成例では、特徴抽出部4からの特徴量(特徴パターン)を通信手段(例えば電話回線や無線など)を介して中央装置32に送信するようにしているので、通信手段(電話回線や無線など)の品質や通信環境が多少悪い場合でも、特徴量(特徴パターン)信号は、音声信号を送信する場合に比べて、影響を受けにくく、従って、中央装置32において、話者認識をより正しく行なうことができる。また、特徴量(特徴



パターン)信号は、音声信号に比べてデータ量がはるかに少なく、特徴量(特徴パターン)信号を中央装置32に送信する場合には、音声信号を送信する場合に比べて、送信時間を著しく短縮することができる。

【0055】さらに、この第2の構成例では、端末側に、音声区間検出部3、特徴抽出部4が設けられていることによって、利用者は、自己の声の特性に適合するよう、音声区間検出部3の特性、特徴抽出部4の特性を管理することができる。例えば、自己の声の音量や音質に合わせて、音声区間検出の感度(声の大きさのしきい値)などを調整したりすることができる。

【0056】また、上述の第1、第2の構成例においては、話者認識部7(話者認識ソフト)については、端末側ではなく、中央装置側に設けられていることによって(すなわち、話者認識部7に関しては中央装置の1ヵ所一括(集中)管理することによって)、例えば、利用者側の端末の性能(例えば端末に搭載されるソフトウェアの性能)によって認識結果が左右されたりするのを回避し、どのような精度、性能の端末が用いられる場合でも、中央装置側において、常に、一定の認識精度で話者認識を行なうことができる。

【0057】また、新たな話者認識方式が開発され、この新たな話者認識方式を導入する場合、中央装置の話者認識部7、すなわち中央装置の話者認識ソフトを入れ替えるだけで、以後全ての認識を新たな認識方式で行なうことができ、この新たな話者認識方式の導入、すなわちバージョンアップを簡単に(容易に)行なうことができる。

【0058】このように、上述の構成例では、仮に端末に話者認識ソフトが搭載されていたとしても、あるいは、話者認識ソフトが搭載されていなくても、中央装置側の話者認識ソフトを利用するため、安定した信頼度の高い話者認識が可能になる。また、中央装置側の話者認識ソフトを入れ替えるだけで、以後全ての認識を新しい話者認識ソフトで行なうことができることから、そのメンテナンス(例えばバージョンアップ)をも容易に行なうことができる。

【0059】なお、第1、第2の構成例では、上述したように、端末側に、音声区間検出部3、特徴抽出部4が設けられていることによって、利用者は、自己の声の特性に適合するよう、音声区間検出部3の特性、特徴抽出部4の特性を管理することができるという利点を有しているが、その反面、音声区間検出、特徴抽出等の処理と話者認識の処理とが、端末側と中央装置側とで分散してなされるため、中央装置側の管理者は、正規の利用者を正規の利用者と認めなかったり、他人を正規の利用者と認識したりする、いわゆる誤認識が発生する場合に、この原因を中央装置側だけで一括管理することができない。例えば、この原因が、端末の音声区間検出、特徴抽出によるものか、中央装置の話者認識によるものかを、

中央装置側だけで判別することができない。

【0060】誤認識を生じさせる原因としては、例えば、話者が発生した言葉の語頭や語尾が弱く、音声区間検出が正常に行なわれず、特徴パターン上で、この部分が欠落していることがあり、話者がこれに気付かず何回言い直しても同じような結果になってしまうという場合があり、このときには、何回やり直しても正しい認識ができない。

【0061】このような問題を回避するため、第1、第2の構成例において、例えば中央装置32(例えば話者認識管理手段10または話者認識部7)から端末、例えば31-1に所定の情報を提供し、端末31-1の音声区間検出部3や特徴抽出部4は、中央装置32から提供された情報に基づいて、音声区間検出や特徴量変換を行なうことも可能である。

【0062】例えば、中央装置32側からは、情報として、例えば音声区間検出の感度指示情報を提供することができる。

【0063】この場合、端末側において、利用者が当初音声を発声し、この音声は音声区間検出されるときに、音声区間と判定された前後に、例えば0.5秒程度のデータを付加し、このデータをも含めた音声区間内の音声信号を、そのまま端末側のファイル(図示せず)に保存する。このようにして、端末側で利用者の音声信号がファイルに保存された後、これを特徴量(特徴パターン)に変換して中央装置に送信し、中央装置の話者認識部7で話者認識を行なわせる。話者認識の結果、利用者本人ではないと判定された場合、中央装置32側では、例えば、話者認識用情報記憶部5に付随情報として記憶されている音声区間検出の感度(声の大きさのしきい値等)を高めて、これを端末側に、情報として提供(送信)し、端末側において、ファイルに保存されている音声信号に対して再度、音声区間検出を行なわせる。また、話者認識用情報記憶部5に付随情報として記憶されている音声区間検出の感度を下げて、これを端末側に、情報として提供(送信)し、端末側において、ファイルに保存されている音声信号に対して再度、音声区間検出を行なわせる。

【0064】このようにして、音声区間検出の感度を高めた場合と下げた場合とで、それぞれ話者認識を行ない、いずれかで正しい話者であることが判定されれば、この利用者を正しい話者であると認識することができる。このように、音声区間検出の感度を高めることによって、例えば話者の声が小さい場合に、音声区間が正しく検出されないという事態(検出漏れを起こすという事態)が生じていたのを、改善することができ、また、音声区間検出の感度を下げることによって、例えば話者の音声の前後に雑音が生じるような場合に、音声区間検出部が実際よりも長い音声を検出してしまうという事態が生じていたのを、改善することができる。

【0065】上述の例では、中央装置32側から端末側



に提供する情報として、音声区間検出の感度を例にとったが、第2の構成例においては、特徴量変換のサンプリング周波数を情報として端末側に提供することもでき、この場合には、端末側では、中央装置32からの情報によって特徴量変換のサンプリング周波数を変えることができる。さらに、中央装置32側から端末側には、上記以外の種々の情報を提供することもできる。

【0066】このように、端末に、中央装置から所定情報を与えることで、話者認識のみならず、必要に応じて、話者認識のための音声区間検出や特徴抽出などを、中央装置側から管理、制御することができる。

【0067】上述の各構成例の説明では、話者認識を行なう場合について述べたが、標準パターンの新規登録や変更、更新についても、端末側から同様に行なうことができる。なお、中央装置側において、話者認識に用いたデータによって、自動的に標準パターンを更新する機能が備わっている場合は、端末側からの操作を行わずとも、中央装置側で、自動的に標準パターンの更新を行なうことができる。

【0068】また、上述した各構成例において、ある端末から中央装置に通信手段(有線あるいは無線)を介して伝送される音声信号あるいは特徴量(特徴パターン)信号を、この端末の利用者以外の他人が例えば通信手段から盗聴し、盗聴した音声信号あるいは特徴量(特徴パターン)信号を、この他人が正規の利用者になりすまして、中央装置に送信して話者認識を行ない、正規の話者用のアプリケーションを利用したりすることも考えられる。

【0069】このような他人による悪用を防止するため、上述の各構成例において、端末から音声信号あるいは特徴量(特徴パターン)信号を、通信手段によって、中央装置に送信する際、端末側において、送信しようとする音声信号あるいは特徴量(特徴パターン)信号を暗号化し、暗号化した音声信号あるいは特徴量(特徴パターン)信号を、通信手段を介して、中央装置に送信するようにすることもできる。

【0070】図8は図3の話者認識システム(第1の構成例の話者認識システム)において、音声入力手段1(あるいは音声区間検出手段2)から出力される音声信号を暗号化して中央装置に送信する話者認識システムの構成例を示す図である。図8を参照すると、この話者認識システムでは、端末において、音声入力手段1(あるいは音声区間検出手段2)から出力される音声信号を暗号化するための暗号化手段60がさらに設けられている。

【0071】また、図9は図8の具体例を示すものであり、図4の話者認識装置に対応したものとなっている。図9を参照すると、この話者認識システムでは、端末、例えば11-1において、暗号化手段60が音声区間検出部3の後段に設けられている。

【0072】また、図8、図9において、中央装置12には、端末の暗号化手段60に与えるための暗号情報を

生成する暗号情報生成手段61と、端末の暗号化手段60によって暗号化されて端末から送信された音声信号を暗号情報生成手段61によって生成された暗号情報に基づいて復号化する復号化手段62とが、さらに設けられている。

【0073】このような構成の話者認識システムでは、中央装置12は、端末、例えば11-1から話者認識の対象となる音声信号が送信されるに先立って、暗号情報生成手段61で暗号情報(例えば、暗号鍵)を生成し、音声信号が送信されようとしている端末11-1に暗号情報を提供する。なお、端末11-1に提供した暗号情報は、この端末11-1からの音声信号(この暗号情報により暗号化された音声信号)に対して復号化処理を施すために、中央装置12において保存される。

【0074】このようにして、中央装置12から端末11-1に暗号情報が提供された後、端末の暗号化手段60では、中央装置12から提供された暗号情報に基づいて音声信号を暗号化して中央装置12に送信する。中央装置12では、この端末からの暗号化された音声信号を受信すると、復号化手段62により、端末11-1に提供されたと同じ暗号情報を用いて、暗号化された音声信号を元の音声信号に復元し、しかる後、この音声信号から特徴量(特徴パターン)を抽出して、話者認識を行なう。

【0075】このように、この構成例では、端末11-1から音声信号を暗号化して、通信手段13-1を介し、中央装置12に送信するので、他人が通信手段13-1からこれを盗聴しても、この他人は暗号情報を知らない限り、これから音声信号を抽出することができず、従って、利用者の音声信号が他人に盗まれて他人によって悪用されるという事態を有効に防止することができる。

【0076】なお、中央装置12において、暗号情報(暗号鍵)は、端末との通信の度に更新(変更)生成されて、端末に提供され、また、復号化手段62で用いられるのが望ましい。これにより、端末からの音声信号は、毎回、異なった暗号情報によって暗号化され、従って、盗聴によって盗まれた信号は、次回以後、中央装置12によって正しく復号化(復元)されず、正しい話者認識ができなくなるので、他人による盗用を確実に防止することができる。

【0077】図10は図6の話者認識システム(第2の構成例の話者認識システム)において、特徴抽出部4から出力される特徴量(特徴パターン)信号を暗号化して中央装置に送信する話者認識システムの構成例を示す図である。図10を参照すると、この話者認識システムでは、端末において、特徴抽出部4から出力される特徴量(特徴パターン)信号を暗号化する暗号化手段70がさらに設けられている。

【0078】また、図11は図10の具体例を示すもの

であり、図 7 の話者認識装置に対応したものとなっている。図 11 を参照すると、この話者認識システムでは、端末、例えば 31-1 において、暗号化手段 70 が特徴抽出部 4 の後段に設けられている。また、中央装置 32 においては、端末の暗号化手段 70 に与えるための暗号情報を生成する暗号情報生成手段 71 と、端末の暗号化手段 70 によって暗号化されて端末から送信された特徴量(特徴パターン)信号を暗号情報生成手段 71 によって生成された暗号情報に基づいて復号化する復号化手段 72 とが、さらに設けられている。

【0079】このような構成の話者認識システムでは、中央装置 32 は、端末、例えば 31-1 から話者認識の対象となる特徴量(特徴パターン)信号が送信されるに先立って、暗号情報生成手段 71 で暗号情報(例えば、暗号鍵)を生成し、特徴量(特徴パターン)信号が送信されようとしている端末 31-1 に暗号情報を提供する。なお、端末 32-1 に提供した暗号情報は、この端末からの特徴量(特徴パターン)信号(この暗号情報により暗号化された特徴量(特徴パターン)信号)に対して復号化処理を施すために、中央装置 32 において保存される。

【0080】このようにして、中央装置 32 から端末 31-1 に暗号情報が提供された後、端末 31-1 の暗号化手段 70 では、中央装置 32 から提供された暗号情報に基づいて特徴量(特徴パターン)信号を暗号化して中央装置 32 に送信する。中央装置 32 では、この端末 31-1 からの暗号化された特徴量(特徴パターン)信号を受信すると、復号化手段 72 により、端末 31-1 に提供されたと同じ暗号情報を用いて、暗号化された特徴量(特徴パターン)信号を元の特徴量(特徴パターン)信号に復元し、しかる後、この特徴量(特徴パターン)信号により、話者認識を行なう。

【0081】このように、この構成例では、端末 31-1 から特徴量(特徴パターン)信号を暗号化して、通信手段 33-1 を介し、中央装置 32 に送信するので、他人が通信手段 33-1 からこれを盗聴しても、この他人は暗号情報を知らない限り、これから特徴量(特徴パターン)信号を抽出することができず、従って、利用者の特徴量(特徴パターン)信号が他人に盗まれて他人によって悪用されるという事態を有効に防止することができる。

【0082】また、中央装置 32 において、暗号情報(暗号鍵)は、端末との通信の度に更新(変更)生成されて、端末に提供され、また、復号化手段 72 で用いられるのが望ましい。これにより、端末からの特徴量(特徴パターン)信号は、毎回、異なった暗号情報によって暗号化され、従って、盗聴によって盗まれた信号は、次回以後、中央装置 32 によって正しく復号化(復元)されず、正しい話者認識ができなくなるので、他人による盗用を確実に防止することができる。

【0083】上述の各構成例では、音声区間抽出部 3 の後に、特徴抽出部 4 が設けられているが、必要に応じ、

音声区間抽出部 3 の前に、特徴抽出部 4 が設けられるように構成することも可能である。

【0084】また、上述の各構成例において、端末 11-1, 31-1 は、中央装置 12, 32 に音声信号あるいは特徴量(特徴パターン)信号を送信する際、音声信号あるいは特徴量(特徴パターン)信号を一旦ファイルに蓄積し、ファイルに蓄積した音声信号あるいは特徴量(特徴パターン)信号を中央装置 12, 32 に送信するようになっていても良い。また、この場合、図 8 乃至図 11 の構成例において、暗号化手段 60, 70 では、ファイルの前段に設けられても良いし、ファイルの後段に設けられても良い。

【0085】また、図 8~図 11 の各構成例において、暗号情報として暗号鍵を用いた暗号化の手法には、すでに実用化されている様々な方法が利用可能である。例えば、図 10, 図 11 の構成例の場合、中央装置 32 からは、暗号鍵として 1 つの数字を送信することができ、この場合、端末では、この数字の値に基づいて特徴量情報の並び替えを行えば良い。簡単な方法としては、例えば、特徴量情報が  $p_1, \dots, p_i, \dots, p_n$  の  $n$  個の数値で構成され、中央装置 32 から暗号鍵として数値  $C$  が送信されてきた場合、端末は特徴量情報を  $C$  番目の数値から順に、 $p_C, p_{C+1}, \dots$  のように送信し、 $p_n$  まで到達したら  $p_1$  に戻り、 $p_{C-1}$  まで  $n$  個の数値を送信するようにより、特徴量情報に対して暗号化を施すことができる。

【0086】より具体的に、例えば、特徴量情報が  $p_1, \dots, p_{16}$  の 16 個の数値で構成され、暗号鍵として中央装置 32 から“5”が送信されてきた場合、端末は 16 個の数値を、 $p_5, p_6, p_7, p_8, p_9, p_{10}, p_{11}, p_{12}, p_{13}, p_{14}, p_{15}, p_{16}, p_1, p_2, p_3, p_4$  のように順番を入れ替えて送信する。

【0087】上記例では、簡単な暗号化の例を示したが、暗号化の手法はこの例に限定されるものではなく、さらに高度な暗号化の手法を利用すれば、盗用に対してより効果的な防護を行なうことが可能となる。

【0088】例えば、公開鍵方式と呼ばれる手法では、2 つの暗号鍵  $C$  と  $C'$  を使い、受信側はこれらの 2 つの鍵を作成し、1 つ(公開鍵)を送信側へ暗号化のために伝送し、他方を手元に復号化のために保管する。これらの 2 つの暗号鍵は、互いに他方の鍵によって暗号化された情報を復号することは可能であるが、自身によって暗号化された情報は復号することは出来ず、また一方の暗号鍵から他方の暗号鍵を作成・類推することは困難であるように構成されている。

【0089】従って、本発明に公開鍵方式を応用するには、中央装置の暗号情報生成手段でこれら 2 つの鍵を生成し、そのうちの 1 つ  $C$  を端末へ送信し、もう 1 つの鍵  $C'$  を復号化手段で復号化に使用するために保管する。端末では音声信号あるいは音声の特徴量信号を中央装置

からの鍵Cによって暗号化し、中央装置へ送信する。中央装置では、保管されている鍵C'を用いて、端末からの暗号化された音声信号あるいは音声の特徴量信号を復号化する。この構成により、盗聴などの行為によって、端末からの暗号化された音声信号あるいは音声の特徴量信号のみならず、中央装置から端末への暗号情報(暗号鍵)が盗まれたとしても、鍵Cは復号化を行うことは出来ないから、中央装置に保管されている鍵C'が盗まれない限り、端末からの暗号化された音声信号あるいは音声の特徴量信号は復元できず、暗号情報(暗号鍵)の送信毎の更新と組み合わせることで、盗用に対してより効果的な防護を行うことが可能になる。

【0090】また、上述の説明では、話者認識の対象となる音声信号あるいは特徴量(特徴パターン)信号に対して暗号化を施すとしたが、標準パターン登録用の音声信号あるいは特徴量(特徴パターン)信号に対しても同様に暗号化を施して中央装置に送信することができる。さらに、話者認識システムが、標準パターンの変更、更新を端末側からの操作によって行なう機能を有している場合には、標準パターンの変更、更新用の音声信号あるいは特徴量(特徴パターン)信号に対しても同様に暗号化を施して中央装置に送信することができる。

【0091】

【発明の効果】以上に説明したように、請求項1乃至請求項10記載の発明によれば、例えば利用者の自宅あるいは会社等に設置されている端末を用いて、銀行などに設置されている中央装置にアクセスして、話者認識を行なわせ、話者認識の結果、本人であることが確認された後、入出金、残高照会などのアプリケーションを利用することができる。また、仮に端末に話者認識ソフトが搭載されていたとしても、あるいは、話者認識ソフトが搭載されていなくても、中央装置側の話者認識ソフトを利用するため、安定した信頼度の高い話者認識が可能になる。また、中央装置側の話者認識ソフトを入れ替えるだけで、以後全ての認識を新しい話者認識ソフトで行なうことができることから、そのメンテナンス(例えばバージョンアップ)をも容易に行なうことができる。

【0092】特に、請求項1、請求項7記載の発明によれば、端末としては、単に、音声入力手段(あるいは、これとともに指定手段)だけが備わったものであっても良く、従って、端末として、通常の電話器、携帯電話を用いることができる。

【0093】また、請求項2、請求項8記載の発明によれば、特徴抽出部からの特徴量(特徴パターン)を通信手段(例えば電話回線や無線など)を介して中央装置に送信するようにしているので、通信手段(電話回線や無線など)の品質や通信環境が多少悪い場合でも、特徴量(特徴パターン)信号は、音声信号を送信する場合に比べて、影響を受けにくく、従って、中央装置において、話者認識をより正しく行なうことができる。また、特徴量(特

徴パターン)信号は、音声信号に比べてデータ量のはるかに少なく、特徴量(特徴パターン)信号を中央装置に送信する場合には、音声信号を送信する場合に比べて、送信時間を著しく短縮することができる。

【0094】また、請求項3、請求項9記載の発明によれば、端末に、中央装置から所定情報を与えることで、話者認識のみならず、必要に応じて、話者認識のための音声区間検出や特徴抽出なども、中央装置側から管理、制御することができる。

【0095】また、請求項4、請求項5、請求項10記載の発明によれば、端末において音声信号あるいは特徴量(特徴パターン)信号を暗号化して中央装置に送信するので、他人が通信手段からこれを盗聴しても、この他人は暗号情報を知らない限り、これから音声信号、特徴量(特徴パターン)信号を抽出することができず、従って、利用者の音声信号、特徴量(特徴パターン)信号が他人に盗まれて他人によって悪用されるという事態を有効に防止することができる。

【0096】また、請求項6記載の発明によれば、端末からの音声信号あるいは特徴量(特徴パターン)信号は、毎回、異なった暗号情報によって暗号化され、従って、盗聴によって盗まれた信号は、次回以後、中央装置によって正しく復号化(復元)されず、正しい話者認識ができなくなるので、他人による盗用を確実に防止することができる。

【図面の簡単な説明】

【図1】一般的な話者認識システムの構成例を示す図である。

【図2】話者認識用情報記憶部の構成例を示す図である。

【図3】本発明に係る話者認識システムの第1の構成例を示す図である。

【図4】図3の話者認識システムの具体例を示す図である。

【図5】話者認識用情報記憶部の構成例を示す図である。

【図6】本発明に係る話者認識システムの第2の構成例を示す図である。

【図7】図6の話者認識システムの具体例を示す図である。

【図8】本発明に係る話者認識システムの他の構成例を示す図である。

【図9】図8の話者認識システムの具体例を示す図である。

【図10】本発明に係る話者認識システムの他の構成例を示す図である。

【図11】図10の話者認識システムの具体例を示す図である。

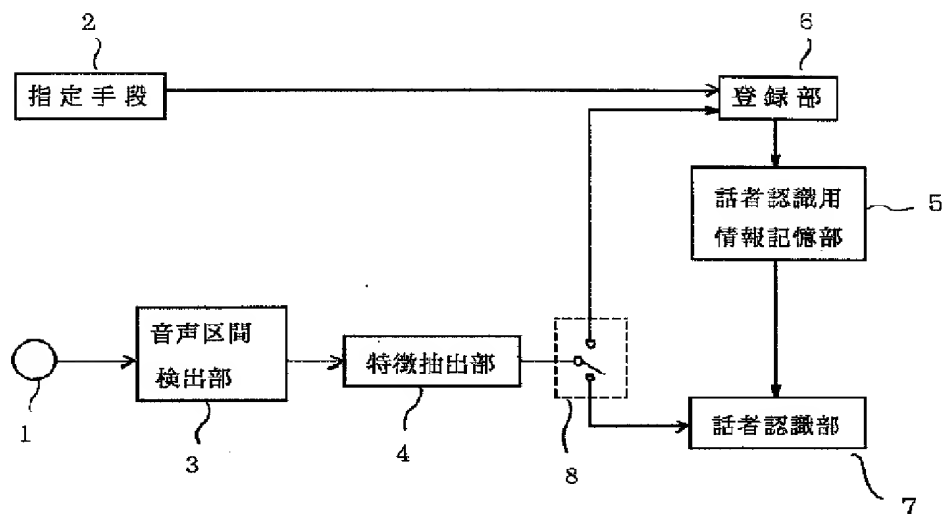
【符号の説明】

1

音声入力手段

2	指示手段	1 0	話者認識管理手段
3	音声区間検出部	1 1, 3 1	端末
4	特徴抽出部	1 2, 3 2	中央装置
5	話者認識用情報記憶部	1 3, 3 3	通信手段
6	登録部	6 0, 7 0	暗号化手段
7	話者認識部	6 1, 7 1	暗号情報生成手段
8	切替部	6 2, 7 2	復号化手段

【図 1】



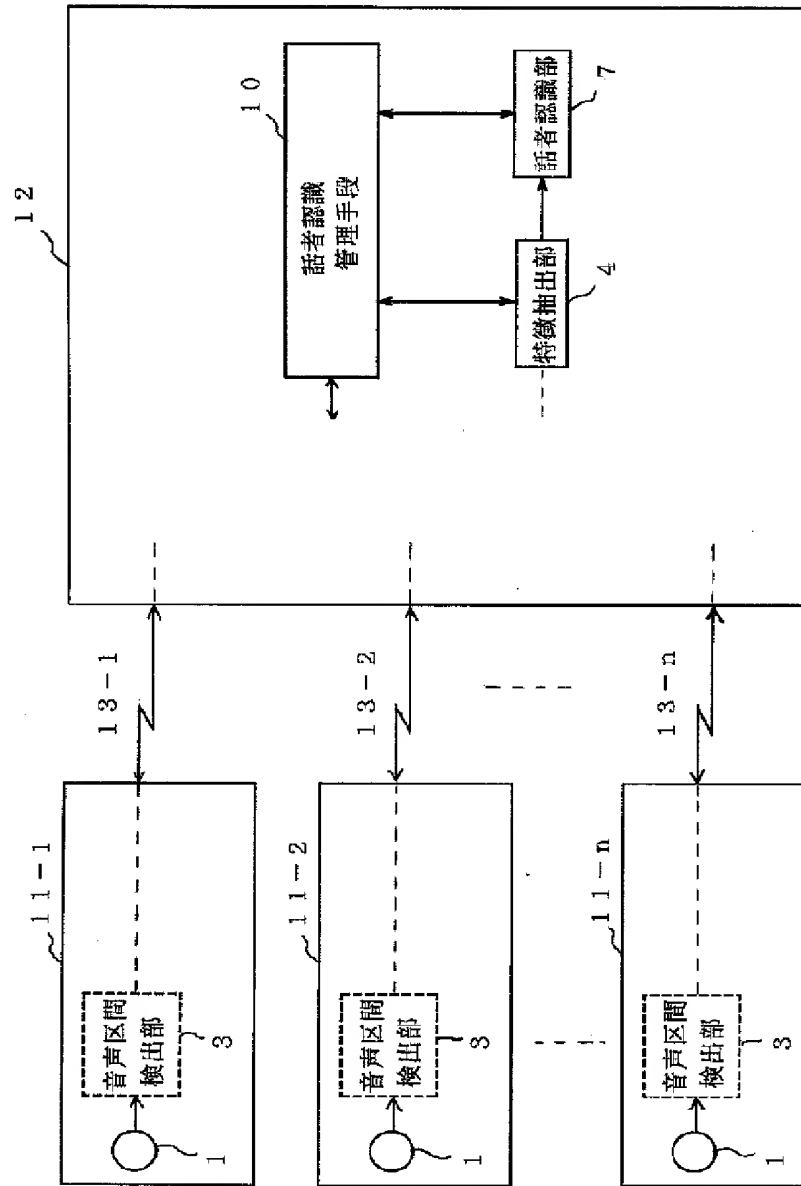
【図 2】

指定情報	標準パターン
(利用者A)の指定情報	(利用者A)の標準パターン
(利用者B)の指定情報	(利用者B)の標準パターン
(利用者C)の指定情報	(利用者C)の標準パターン
(利用者D)の指定情報	(利用者D)の標準パターン
⋮	⋮

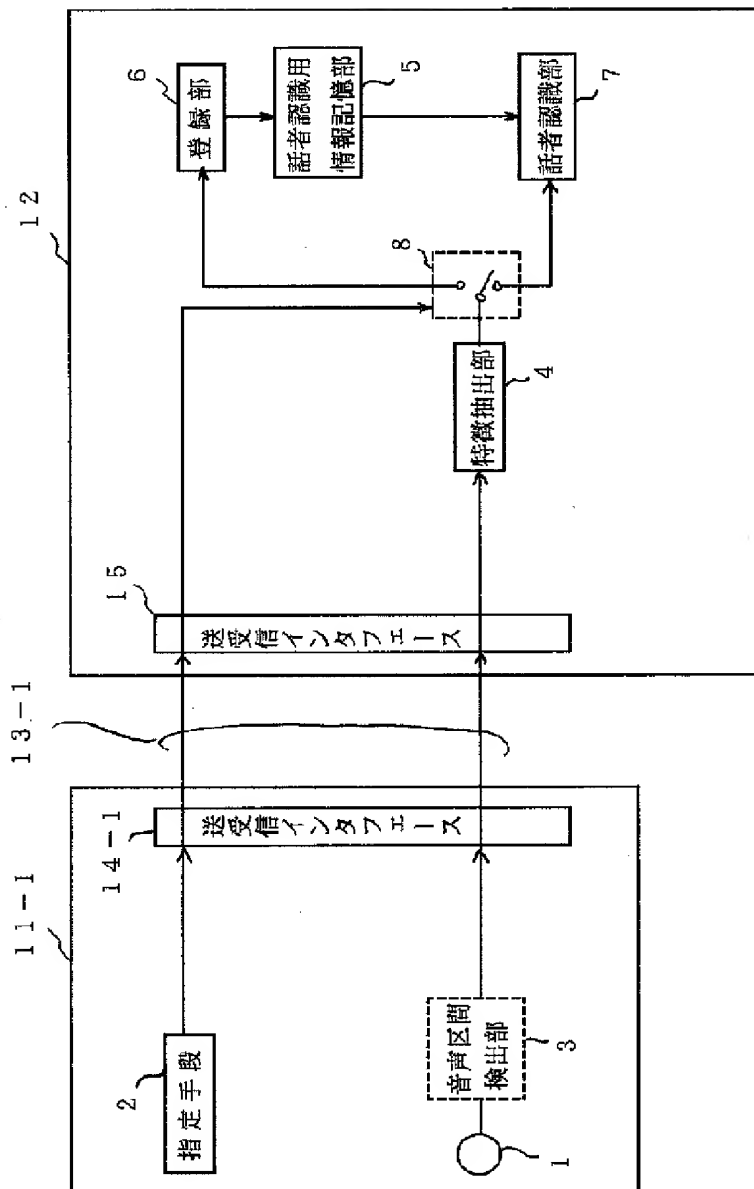
【図 5】

指定情報	標準パターン	付随情報		
		更新日時	しきい値	---
(利用者A)の指定情報	(利用者A)の標準パターン		$TH_A$	
(利用者B)の指定情報	(利用者B)の標準パターン		$TH_B$	
(利用者C)の指定情報	(利用者C)の標準パターン		$TH_C$	
(利用者D)の指定情報	(利用者D)の標準パターン		$TH_D$	
⋮	⋮		⋮	

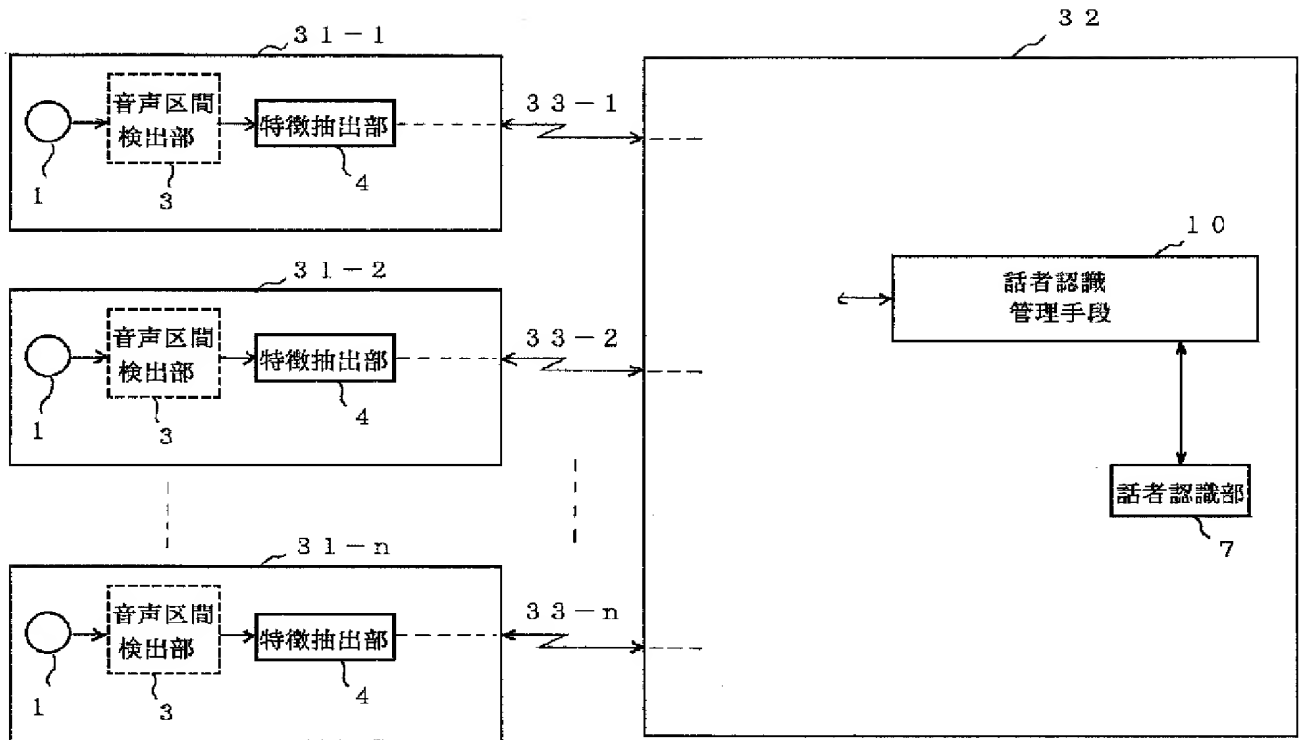
【図3】



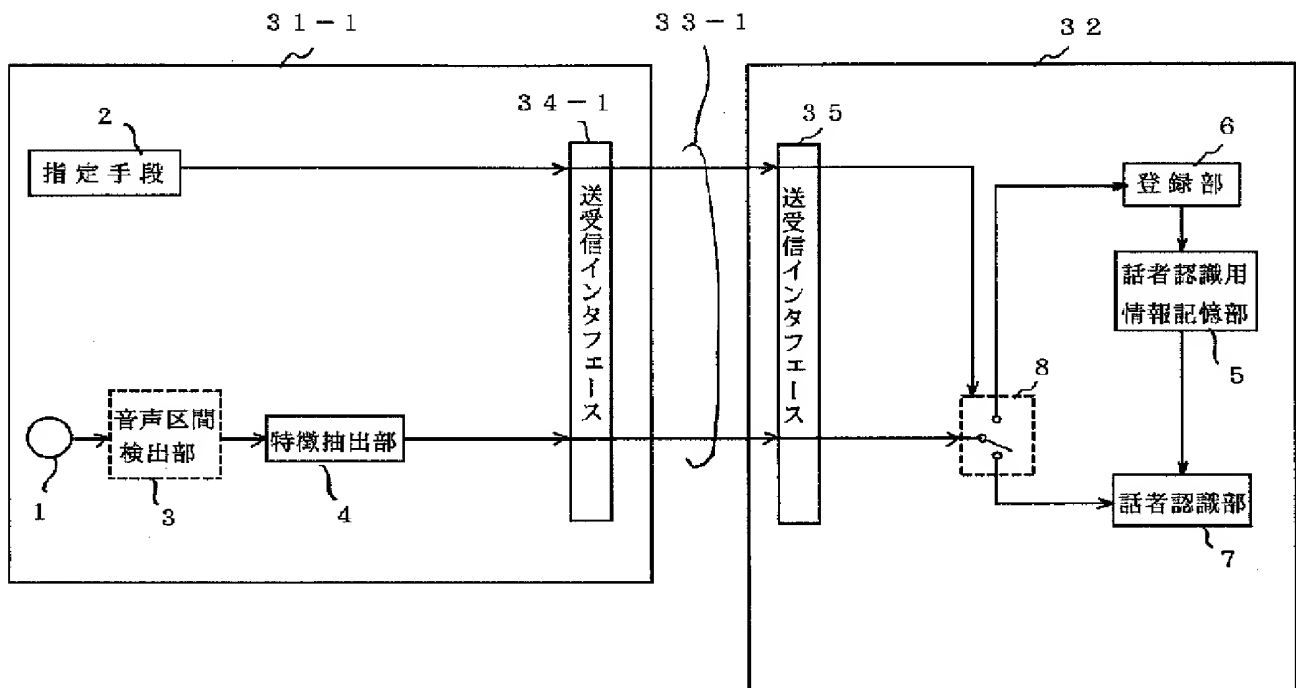
【図4】



【図6】

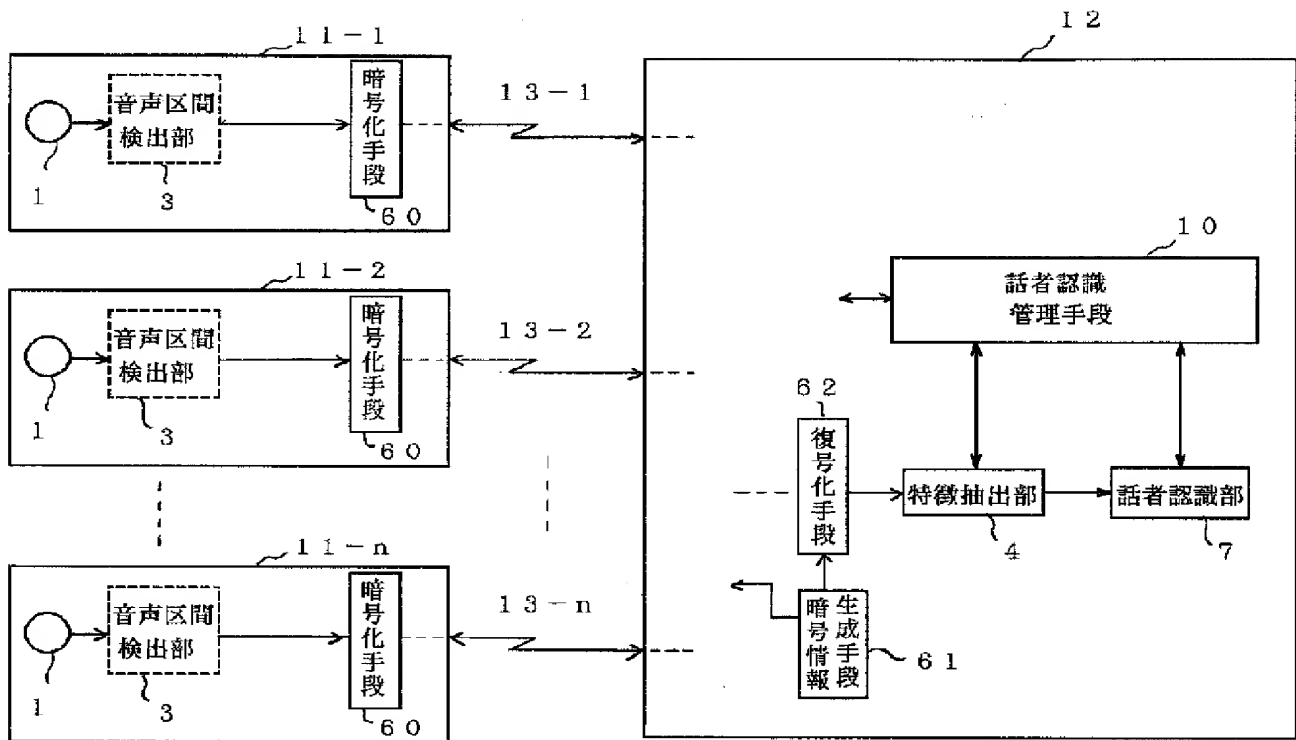


【図7】

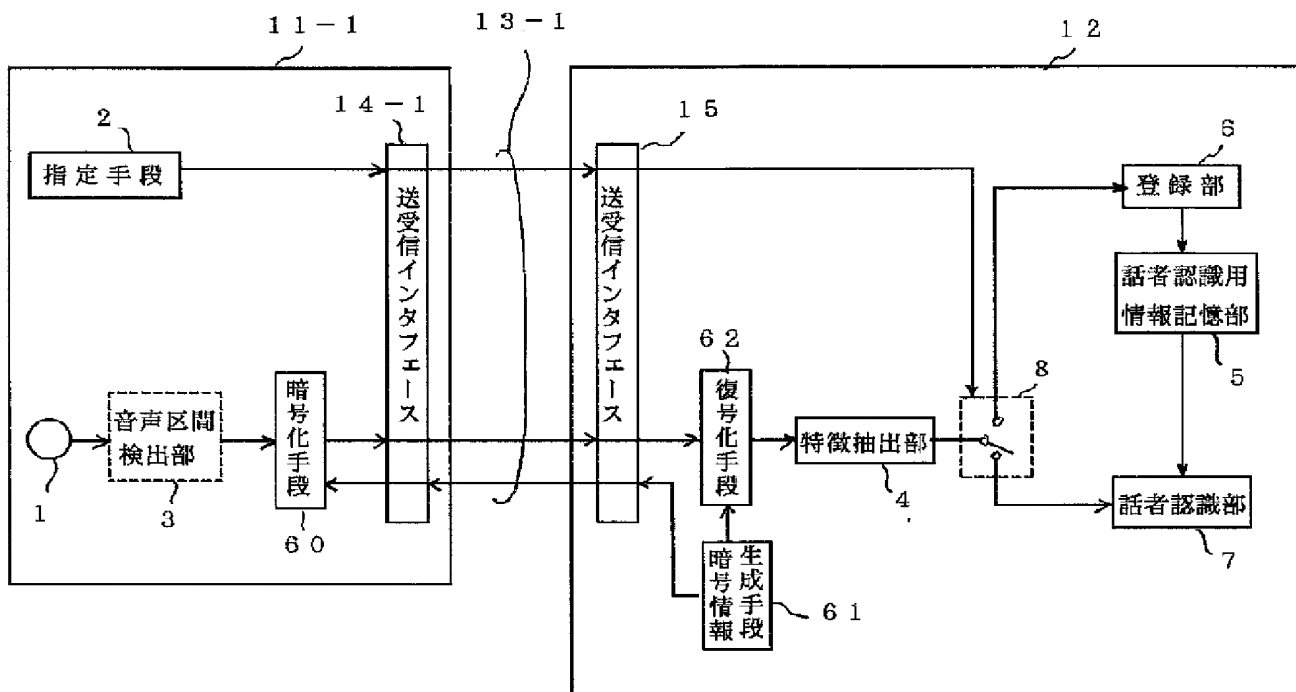




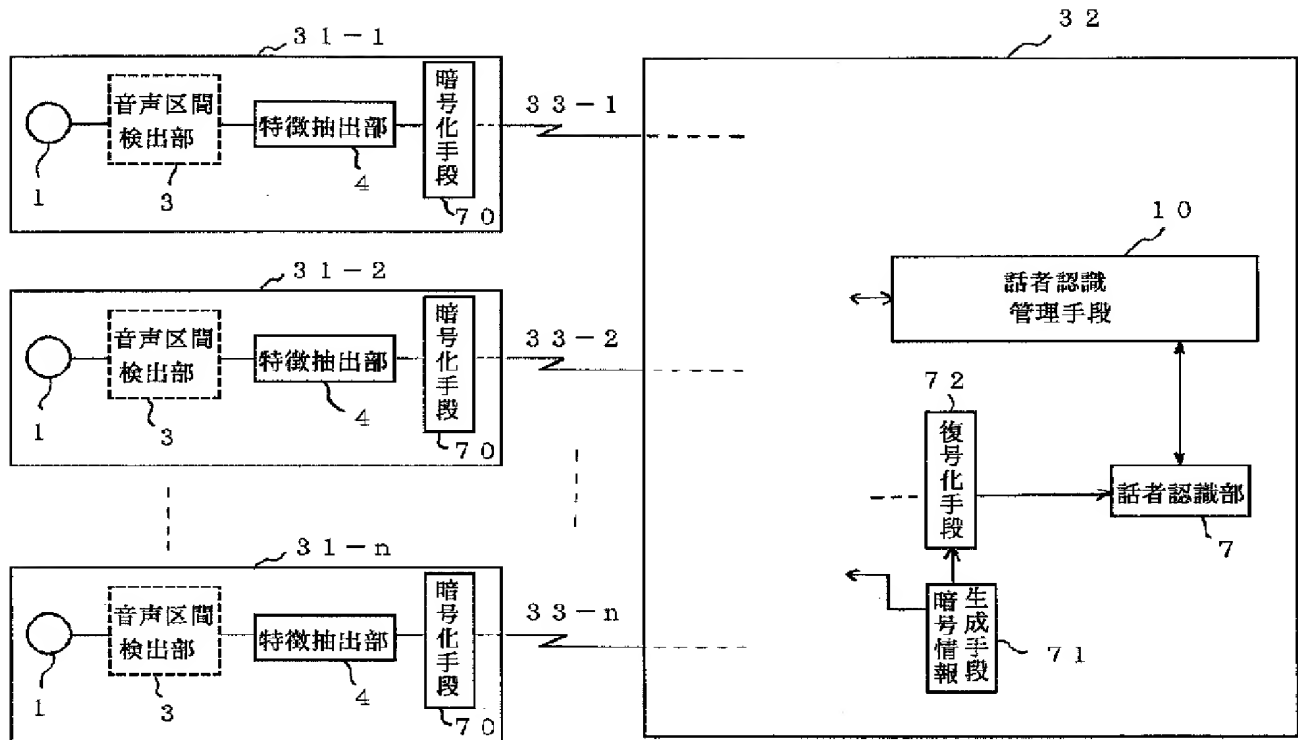
【図8】



【図9】



【図10】



【図11】

